UNITED STATES PATENT APPLICATION

MEMORY CONFIGURATION APPARATUS, SYSTEMS, AND METHODS

INVENTOR

Inching Chen

Schwegman, Lundberg, Woessner & Kluth, P.A.
1600 TCF Tower
121 South Eighth Street
Minneapolis, MN 55402
ATTORNEY DOCKET SLWK 884.B97US1
Client Reference P18357

MEMORY CONFIGURATION APPARATUS, SYSTEMS, AND METHODS

Technical Field

[0001] Various embodiments described herein relate to information processing generally, such as apparatus, systems, and methods used to store and retrieve data.

Background Information

[0002] Computer system performance may be highly dependent on associated memory system operational efficiency. For example, processing that stalls when data is unavailable can render results at an unacceptably slow rate.

[0003] Many memory systems have been designed to support a fixed data bus width, with an implicit fixed data processing bandwidth. However, different applications may be able to use a variety of memory access data widths, such as 32 bits, 64 bits, and 128 bits, among others. As a result, there may be a disparity between the number of data bits available from memory and the number of data bits sought by the processor when executing a particular application. If the available bandwidth is too small, extra memory cycles may be required to process a given amount of data. If the available bandwidth is too large, the existing memory capacity may not be used efficiently.

Brief Description of the Drawings

[0004] FIG. 1 is a memory address-mapping table according to various embodiments;

[0005] FIG. 2 is a block diagram of an apparatus and a system according to various embodiments;

[0006] FIG. 3 is a flow chart illustrating several methods according to various embodiments; and

[0007] FIG. 4 is a block diagram of an article according to various embodiments.

Detailed Description

[0008] In some embodiments, a memory subsystem may have a bandwidth scalable according to various received indications, such as a particular hardware or software operating configuration, or the number of processors to be used in carrying out certain kinds of operations. For example, in a reconfigurable communications architecture (RCA), the number of data processing units (DPUs) used in a processing element (PE) may change according to the processing requirements of a particular function (e.g., single-instruction, multiple-data (SIMD) and multiple-instruction, multiple-data (MIMD) operations) and/or application.

[0009] In some embodiments, memory address access mapping may operate to support both MIMD and SIMD types of operation. For example, assuming that a memory subsystem includes four banks of interleaved memory, and depending on the number of DPUs used, one, two, or four banks of memory may be substantially simultaneously deployed for data storage and retrieval. Thus, in some embodiments, the memory subsystem associated with an RCA may operate to change bandwidth (e.g., the number of bits accessed in a memory read and/or write operation) based on the number of processors in use.

[0010] For the purposes of this document, the term "energy conduit" includes any type of device or apparatus that has the capability to transmit and/or receive energy to and/or from space. Examples of such energy conduits include antennas, infra-red transmitters, infra-red receivers, photo-emitters (e.g., light emitting diodes), photo-receptors (e.g., a photocell), and charge-coupled devices, among others.

[0011] A "protocol type" includes any rule or guide for formatting data as it is communicated, with respect to time, space, frequency, and/or some other organizing function. Examples of such formatting mechanisms include communications standards (e.g., IEEE 802.11, Bluetooth), communications

protocols (e.g., transmission control protocol/Internet protocol), spread-spectrum communications techniques, multiple carrier communications techniques (e.g., orthogonal frequency division multiplexing), a data format (e.g., time division, multiple access), and various data types (e.g., binary, alphanumeric, audio, video). Other protocol types may include an operation type, such as MIMD and SIMD operation types, and instruction types, such as various classes or groups of instructions types that may be executed by a microprocessor. For more information regarding some of the formatting mechanisms mentioned above, please refer to "IEEE Standards for Information Technology -- Telecommunications and Information Exchange between Systems -- Local and Metropolitan Area Network -- Specific Requirements -- Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY), ISO/IEC 8802-11: 1999" and "Bluetooth System Specification, Bluetooth Special Interest Group, Ver. 1.1, March 2001", and related amendments.

[0012] The term "transceiver" (e.g., a device including a transmitter and a receiver) may be used in place of either "transmitter" or "receiver" throughout this document. Thus, anywhere the term transceiver is used, "transmitter" and/or "receiver" may be substituted.

[0013] FIG. 1 is a memory address-mapping table 100 according to various embodiments. In this figure, a specific implementation of some embodiments is shown for simplicity, and it should not be used to limit the embodiments disclosed. Thus, assuming that a memory subsystem includes four banks of interleaved memory, and, depending on the number of DPUs used, that one, two, or four banks of memory can be substantially simultaneously deployed for data storage and retrieval, as noted above, some possible examples of address mapping required for supporting MIMD and SIMD operations are given in the table 100.

[0014] In some embodiments, as an example of MIMD operations, all DPUs may read memory location 0, but the actual data can come from locations [0], [1], [2], or [3], respectively (e.g., in 32-bit mode). In an example of SIMD operations,

reading memory location 0 can result in reading data from locations [0 & 1] or [2 & 3] (e.g., in 64-bit mode), or [0, 1, 2 & 3] (e.g., in 128-bit mode).

[0015] Calculation of the physical mapped addresses to each memory bank for various modes of memory access (e.g., 32-bit word access, 64-bit two-word access, or 128-bit four-word access) can thus be simplified. A generalized mapping formula can be given as:

$$ADDR = (M/ACC\_BANKS)*TOT\_BANKS+OFFSET$$

where ADDR is the physical address to be accessed, M is the logical memory address in 32-bit words, ACC_BANKS is the number of banks to be accessed at a time (e.g., a group size of one bank for 32-bit mode, two banks for 64-bit mode, etc.), TOT_BANKS is the total number of banks that can be accessed (e.g., four banks in the assumed example), and OFFSET is the bank to be accessed.

[0016] For example, assume the existence of 8,192 bytes of memory having logical memory addresses having 13 bits M. Then the physical mapping address ADDR may be generated as a one or two-bit shift of M to the left (for 64 and 128-bit access, respectively), and an offset of 0, 1, 2, or 3 AND'ed with the shifted value of M. Thus, the hardware for an address generation unit in some embodiments may make use of a two-bit left/right shifter and/or an adder.

[0017] FIG. 2 is a block diagram of an apparatus 200 and a system 210 according to various embodiments, each of which may operate in the manner described above. For example, an apparatus 200 may comprise an RCA 212 including one or more selection modules 214 to select a memory access group size of about 2^N memory banks 218 (e.g., 1, 2, or 4 banks) for one or more DPUs SU0, SU1, SU2, and SU3. The memory banks 218 may be multi-port memory banks. Some of the ports may be accessed internally using the DPUs SU0, SU1, SU2, and SU3, and some of the ports may be accessed externally, using other processing devices. The selection module 214, which can include a crossbar, may respond to receiving an indication 222 of a change in a protocol type, such as receiving an

indication 222 of a change in an operation type (e.g., a SIMD operation, a MIMD operation, and a combination of SIMD and MIMD operations), and/or an instruction type. For example, such indications 222 may be received in the form of configuration instructions, such as operation codes. The group of banks 218 may be selected from a number B of banks. N may be associated with the protocol type and selected so that the memory access group size 2^N is less than or equal to B. For example, in some embodiments, the memory access group size 2^N may be selected from a group of numbers including two raised to a positive integer power, such as 2, 4, ..., B.

[0018] One or more of the DPUs SU0, SU1, SU2, and SU3 may be capable of addressing the number B of banks, and some or all of the DPUs SU0, SU1, SU2, and SU3 may be included in a single processing element (PE) 226.

[0019] Any number of PEs 228, including input/output (I/O) PEs, may be included in an RCA 212. Each of the PEs 228 may include one or more system controllers 229. The apparatus 200 may include a hardware or software address generator 230 to generate a physical address located in the 2^N memory banks 218. Thus, in some embodiments, the apparatus 200 may include a hardware element 234 to store a plurality of output indications 238 based on a corresponding plurality of memory access group sizes, which may in turn be selected responsive to a corresponding plurality of protocol type indications 222.

[0020] Thus, in some embodiments, an apparatus 200 may include, in combination, a number B of memory banks addressable using a memory access group size of about 2^N memory banks 218 responsive to receiving an indication 222 of a change in a protocol type, including an operation type, as well as an instruction type in order to map a plurality of logical addresses associated with the 2^N memory banks 218 to a plurality of physical addresses associated with the number B of banks . The group of banks 218 may be selected from the number B of banks, wherein N may be associated with the protocol type, and wherein the memory access group size 2^N may be less than or equal to about B. The protocol type may include a SIMD operation type, a MIMD operation type, and a combination of the

SIMD and MIMD operation types. In many embodiments, the memory access group size may be reprogrammable and selectable in software.

[0021] In some embodiments, the processing element 226 may comprise an RCA processing element having four DPUs SU0-SU3 for fixed-point arithmetic processing (e.g., 16-bit Multiply and Add), four banks of memory 218, and a control unit (CU) 242. The DPUs SU0-SU3 may operate independently (e.g., in carrying out a MIMD operation), and in tandem (e.g., in carrying out a SIMD operation). The CU 242 may operate to control memory bank 218 access and crossbar configuration, provide operation codes to one or more of the processors SU0-SU3, and load/store data for the processors SU0-SU3. In some embodiments, multiple CUs 242 may be included in a single PE 226, 228. In this case, each CU 242 may be associated with a memory access group size or a selected group of DPUs SU0-SU3 (e.g., 1, 2, or four DPUs SU0-SU3). Other embodiments may be realized.

[0022] For example, a system 210 may include an apparatus 200, as well as an energy conduit 250 to transmit data 254 processed by one or more of the DPUs SU0-SU3. In some embodiments, the energy conduit may be selected from one or more of an antenna, an infra-red transmitter, an infra-red receiver, a photo-emitter, a photo-receptor, and/or a charge-coupled device. In some embodiments, the system 210 may include a bus 258 to couple one of the DPUs SU0-SU3 to each other and/or to the banks 218, including one of the number B of banks in a selected group.

[0023] In some embodiments, the system 210 may include a memory 262 to store a plurality of memory access group sizes indexed to a corresponding plurality of protocol types, such as operation types, including instruction types. In some embodiments, the system 210 may include a transceiver 266 to couple one or more of PEs 228, including one or more DPUs SU0-SU3 to the energy conduit 250.

[0024] The apparatus 200, system 210, RCA 212, selection module 214, memory banks 218, indication 222, processing elements 226, 228, system controller 229, address generator 230, hardware element 234, output indication 238, CU 242, data 254, bus 258, transceiver 266, and DPUs SU0, SU1, SU2, and SU3 may all be characterized as "modules" herein. Such modules may include hardware circuitry,

and/or one or more processors and/or memory circuits, software program modules, including objects and collections of objects, and/or firmware, and combinations thereof, as desired by the architect of the apparatus 200 and the system 210, and as appropriate for particular implementations of various embodiments.

[0025] It should also be understood that the apparatus and systems of various embodiments can be used in applications other than transmitters and receivers, and other than for wireless systems, and thus, various embodiments are not to be so limited. The illustrations of an apparatus 200 and system 210 are intended to provide a general understanding of the structure of various embodiments, and they are not intended to serve as a complete description of all the elements and features of apparatus and systems that might make use of the structures described herein.

[0026] Applications that may include the novel apparatus and systems of various embodiments include electronic circuitry used in high-speed computers, communication and signal processing circuitry, modems, processor modules, embedded processors, data switches, and application-specific modules, including multilayer, multi-chip modules. Such apparatus and systems may further be included as sub-components within a variety of electronic systems, such as televisions, cellular telephones, personal computers, personal digital assistants (PDAs), workstations, radios, video players, vehicles, and others.

[0027] FIG. 3 is a flow chart illustrating several methods according to various embodiments. In some embodiments of the invention, a method 311 may begin with selecting a memory access group size of about 2^N memory banks responsive to receiving an indication of a change in a protocol type, such as an operation type, including an instruction type at block 321. This activity may be accomplished in order to map a plurality of logical addresses associated with the 2^N memory banks to a plurality of physical addresses associated with the number B of banks.

[0028] As noted previously, the group of banks may be selected from a number B of banks. N may be associated with the protocol type, including an

operation type, and may be selected so that 2^N is less than or equal to about B. Indications of a protocol type may be selected from hardware indications and software indications, among others. The protocol types may include one or more operation types, including SIMD and MIMD operation types. Other embodiments may be realized.

[0029] For example, in some embodiments, the method 311 may include selecting a first group size for a first DPU at block 321, which is different than a second group size selected for a second DPU at block 331. Each of the DPUs may be capable of addressing the number B of banks, and the sum total of memory banks accessed by all of the DPUs taken together may be limited to the number B of banks. In some embodiments, the memory access group size may be associated with a selected number of access bits (e.g., 32-bits, 64-bits, etc. as described previously).

[0030] In some embodiments, the method 311 may include configuring a crossbar at block 341 to operate using a memory access group size responsive to receiving an indication of a change in the protocol type, such as from one wireless protocol to another, or from one operation type (e.g., SIMD) to another (e.g., MIMD). Many other embodiments may be realized.

[0031] For example, a method 361 may include controlling a bandwidth of a memory coupled to a plurality of DPUs responsive to the number of DPUs in use at block 371 so that a plurality of logical addresses associated with the memory is mapped to a plurality of physical addresses associated with the memory. The number of DPUs in use may in turn be responsive to an indication provided by an application to be executed. And, as noted above, the bandwidth of the memory may be associated with a selected number of access bits provided by the plurality of data processing units. In some embodiments, controlling the bandwidth at block 371 may further include controlling an address-mapping function of the memory at block 381.

[0032] It should be noted that the methods described herein do not have to be executed in the order described, or in any particular order. Moreover, various

activities described with respect to the methods identified herein can be executed in serial or parallel fashion. For the purposes of this document, the terms "information" and "data" may be used interchangeably. Information, including parameters, commands, operands, and other data, can be sent and received in the form of one or more carrier waves.

[0033] Upon reading and comprehending the content of this disclosure, one of ordinary skill in the art will understand the manner in which a software program can be launched from a computer-readable medium in a computer-based system to execute the functions defined in the software program. One of ordinary skill in the art will further understand the various programming languages that may be employed to create one or more software programs designed to implement and perform the methods disclosed herein. The programs may be structured in an object-orientated format using an object-oriented language such as Java, Smalltalk, or C++. Alternatively, the programs can be structured in a procedure-orientated format using a procedural language, such as assembly or C. The software components may communicate using any of a number of mechanisms well-known to those skilled in the art, such as application program interfaces or inter-process communication techniques, including remote procedure calls. The teachings of various embodiments are not limited to any particular programming language or environment, including Hypertext Markup Language (HTML) and Extensible Markup Language (XML). Thus, other embodiments may be realized, as shown in FIG. 4.

[0034] FIG. 4 is a block diagram of an article 485 according to various embodiments, such as a computer, a memory system, a magnetic or optical disk, some other storage device, and/or any type of electronic device or system. The article 485 may comprise a processor 487 coupled to a machine-accessible medium such as a memory 489 (e.g., a memory including an electrical, optical, or electromagnetic conductor) having associated information 491 (e.g., computer program instructions, and/or other data), which when accessed, results in a machine (e.g., the processor 487) performing such actions as selecting a memory access

group size of about 2^N memory banks responsive to receiving an indication of a change in a protocol type, wherein the group may be selected from a number B of banks, so that a plurality of logical addresses associated with the 2^N memory banks is mapped to a plurality of physical addresses associated with the number B of banks. N may be associated with the protocol type and selected so that 2^N is less than or equal to about B. The protocol type may be selected from at least one of a SIMD operation type, a MIMD operation type, and a combination of SIMD and MIMD operation types. Other activities may include selecting a first memory access group size for a first DPU different than a second memory access group size, and selecting the second memory access group size for a second DPU, wherein the first DPU and the second DPU are capable of addressing the number B of banks.

[0035] Implementing the apparatus, systems, and methods described herein may result in providing a reconfigurable memory in which the number of substantially simultaneously accessible data bits can vary according to an indication given by an application, among others. Scalable memory bandwidth may then be available as needed, and, in some embodiments, unnecessary processing elements and/or memory units may be disabled to save power.

[0036] The accompanying drawings that form a part hereof show by way of illustration, and not of limitation, specific embodiments in which the subject matter may be practiced. The embodiments illustrated are described in sufficient detail to enable those skilled in the art to practice the teachings disclosed herein. Other embodiments may be utilized and derived therefrom, such that structural and logical substitutions and changes may be made without departing from the scope of this disclosure. This Detailed Description, therefore, is not to be taken in a limiting sense, and the scope of various embodiments is defined only by the appended claims, along with the full range of equivalents to which such claims are entitled.

[0037] Thus, although specific embodiments have been illustrated and described herein, it should be appreciated that any arrangement calculated to achieve the same purpose may be substituted for the specific embodiments shown. This disclosure is intended to cover any and all adaptations or variations of various

embodiments. Combinations of the above embodiments, and other embodiments not specifically described herein, will be apparent to those of skill in the art upon reviewing the above description.

[0038] The Abstract of the Disclosure is provided to comply with 37 C.F.R. §1.72(b), requiring an abstract that will allow the reader to quickly ascertain the nature of the technical disclosure. It is submitted with the understanding that it will not be used to interpret or limit the scope or meaning of the claims. In addition, in the foregoing Detailed Description, it can be seen that various features are grouped together in a single embodiment for the purpose of streamlining the disclosure. This method of disclosure is not to be interpreted as reflecting an intention that the claimed embodiments require more features than are expressly recited in each claim. Rather, as the following claims reflect, inventive subject matter lies in less than all features of a single disclosed embodiment. Thus the following claims are hereby incorporated into the Detailed Description, with each claim standing on its own as a separate embodiment.